
Subject: [PATCH 2.6.25] multiple namespaces in the all_dst_ifdown routines

Posted by [den](#) on Thu, 06 Dec 2007 12:17:19 GMT

[View Forum Message](#) <> [Reply to Message](#)

move dst entries to a namespace loopback to catch refcounting leaks.

Signed-off-by: Denis V. Lunev <den@openvz.org>

```
net/core/dst.c      | 4 ++--
net/ipv4/route.c    | 5 +++--
net/ipv4/xfrm4_policy.c | 3 +-
net/ipv6/route.c    | 7 +++++--
net/ipv6/xfrm6_policy.c | 3 +-
net/xfrm/xfrm_policy.c | 2 +-
6 files changed, 15 insertions(+), 9 deletions(-)
```

diff --git a/net/core/dst.c b/net/core/dst.c

index f538061..5c6cfc4 100644

--- a/net/core/dst.c

+++ b/net/core/dst.c

```
@@ -279,11 +279,11 @@ static inline void dst_ifdown(struct dst_entry *dst, struct net_device
*dev,
    if (!unregister) {
        dst->input = dst->output = dst_discard;
    } else {
- dst->dev = init_net.loopback_dev;
+ dst->dev = dst->dev->nd_net->loopback_dev;
    dev_hold(dst->dev);
    dev_put(dev);
    if (dst->neighbour && dst->neighbour->dev == dev) {
- dst->neighbour->dev = init_net.loopback_dev;
+ dst->neighbour->dev = dst->dev;
    dev_put(dev);
    dev_hold(dst->neighbour->dev);
    }
}
```

diff --git a/net/ipv4/route.c b/net/ipv4/route.c

index dae1290..e4aa97e 100644

--- a/net/ipv4/route.c

+++ b/net/ipv4/route.c

```
@@ -1425,8 +1425,9 @@ static void ipv4_dst_ifdown(struct dst_entry *dst, struct net_device
*dev,
{
    struct rtable *rt = (struct rtable *) dst;
    struct in_device *idev = rt->idev;
- if (dev != init_net.loopback_dev && idev && idev->dev == dev) {
- struct in_device *loopback_idev = in_dev_get(init_net.loopback_dev);
+ if (dev != dev->nd_net->loopback_dev && idev && idev->dev == dev) {
+ struct in_device *loopback_idev =
```

```

+ in_dev_get(dev->nd_net->loopback_dev);
  if (loopback_iddev) {
    rt->idev = loopback_iddev;
    in_dev_put(iddev);
diff --git a/net/ipv4/xfrm4_policy.c b/net/ipv4/xfrm4_policy.c
index 19fdf8a..e086260 100644
--- a/net/ipv4/xfrm4_policy.c
+++ b/net/ipv4/xfrm4_policy.c
@@ -216,7 +216,8 @@ static void xfrm4_dst_ifdown(struct dst_entry *dst, struct net_device
 *dev,

    xdst = (struct xfrm_dst *)dst;
    if (xdst->u.rt.idev->dev == dev) {
- struct in_device *loopback_iddev = in_dev_get(init_net.loopback_dev);
+ struct in_device *loopback_iddev =
+ in_dev_get(dev->nd_net->loopback_dev);
    BUG_ON(!loopback_iddev);

    do {
diff --git a/net/ipv6/route.c b/net/ipv6/route.c
index e36cac9..e757a3c 100644
--- a/net/ipv6/route.c
+++ b/net/ipv6/route.c
@@ -216,9 +216,12 @@ static void ip6_dst_ifdown(struct dst_entry *dst, struct net_device *dev,
 {
    struct rt6_info *rt = (struct rt6_info *)dst;
    struct inet6_dev *idev = rt->rt6i_iddev;
+ struct net_device *loopback_dev =
+ dev->nd_net->loopback_dev;

- if (dev != init_net.loopback_dev && idev != NULL && idev->dev == dev) {
- struct inet6_dev *loopback_iddev = in6_dev_get(init_net.loopback_dev);
+ if (dev != loopback_dev && idev != NULL && idev->dev == dev) {
+ struct inet6_dev *loopback_iddev =
+ in6_dev_get(loopback_dev);
    if (loopback_iddev != NULL) {
      rt->rt6i_iddev = loopback_iddev;
      in6_dev_put(iddev);
diff --git a/net/ipv6/xfrm6_policy.c b/net/ipv6/xfrm6_policy.c
index cc0d151..7b360ea 100644
--- a/net/ipv6/xfrm6_policy.c
+++ b/net/ipv6/xfrm6_policy.c
@@ -233,7 +233,8 @@ static void xfrm6_dst_ifdown(struct dst_entry *dst, struct net_device
 *dev,

    xdst = (struct xfrm_dst *)dst;
    if (xdst->u.rt6.rt6i_iddev->dev == dev) {
- struct inet6_dev *loopback_iddev = in6_dev_get(init_net.loopback_dev);

```

```
+ struct inet6_dev *loopback_idev =
+ in6_dev_get(dev->nd_net->loopback_dev);
  BUG_ON(!loopback_idev);

  do {
diff --git a/net/xfrm/xfrm_policy.c b/net/xfrm/xfrm_policy.c
index a9ac748..900f6b6 100644
--- a/net/xfrm/xfrm_policy.c
+++ b/net/xfrm/xfrm_policy.c
@@ -1932,7 +1932,7 @@ static int stale_bundle(struct dst_entry *dst)
void xfrm_dst_ifdown(struct dst_entry *dst, struct net_device *dev)
{
  while ((dst = dst->child) && dst->xfrm && dst->dev == dev) {
- dst->dev = init_net.loopback_dev;
+ dst->dev = dev->nd_net->loopback_dev;
  dev_hold(dst->dev);
  dev_put(dev);
}
--
1.5.3.rc5
```

Subject: Re: [PATCH 2.6.25] multiple namespaces in the all dst_ifdown routines
Posted by [davem](#) on Fri, 07 Dec 2007 08:39:14 GMT
[View Forum Message](#) <> [Reply to Message](#)

From: "Denis V. Lunev" <den@openvz.org>
Date: Thu, 6 Dec 2007 15:17:46 +0300

> move dst entries to a namespace loopback to catch refcounting leaks.
>
> Signed-off-by: Denis V. Lunev <den@openvz.org>

Applied, thanks.
